Single Top Wt channel: b-tagging Studies Update

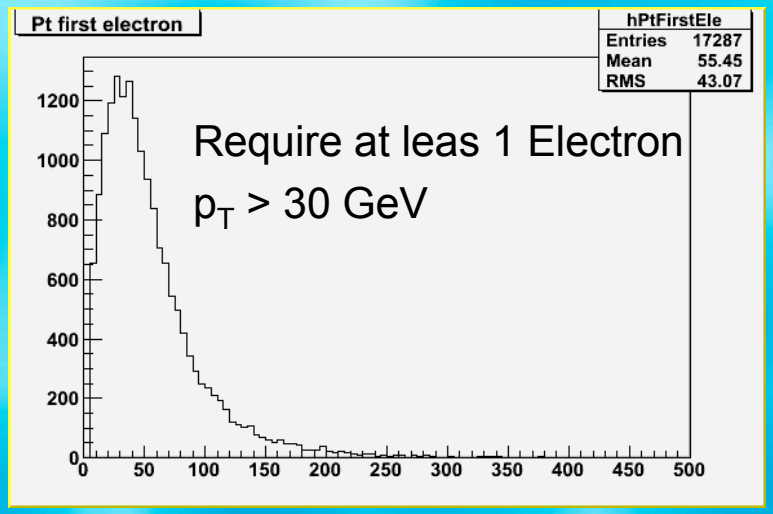
S. Rolli
Tufts University
M. Cobal, M.P. Giordani,
University of Udine

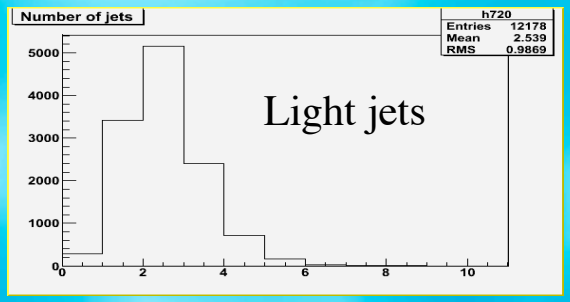
Summary of previous activities

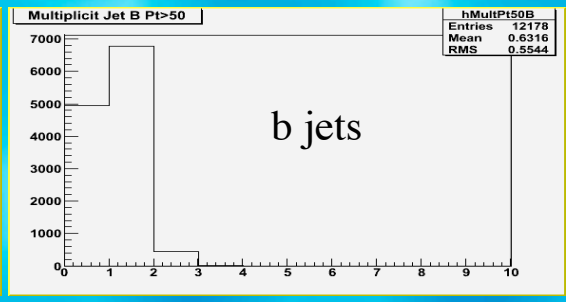
- CBNT and AOD preliminary studies performed for Rome workshop:
  - Starting point was to reproduce the TDR numbers;
  - Final goal is to complete the analysis with full simulation, all background sources and new analysis tools.

CNBT Studies Summary

40000 Wt generated with TopRex
rome.004530.evgen.wt_ph_ml._0000X.pool.root
X=1,9 ($W^- \rightarrow l^- \upsilon$ $W^+ \rightarrow jj$)
Standard Atlfast run on it, relevant parameters:
Electrons: pt > 5 GeV, $|\eta|<2.5$
Jets: Cone 0.4, pT > 5 GeV



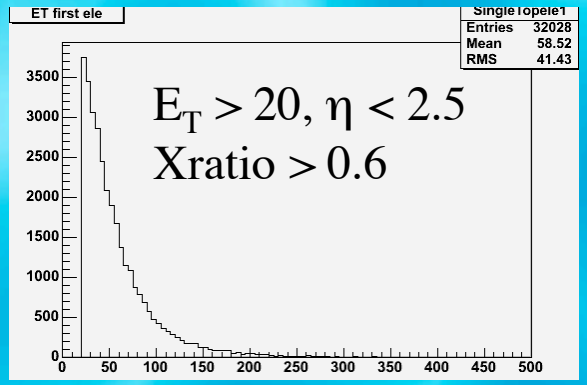




All evts	40000
1 lepton	12178
1 b jets pt 50	6788
2 light jet pt 30	2873 (7.1%)

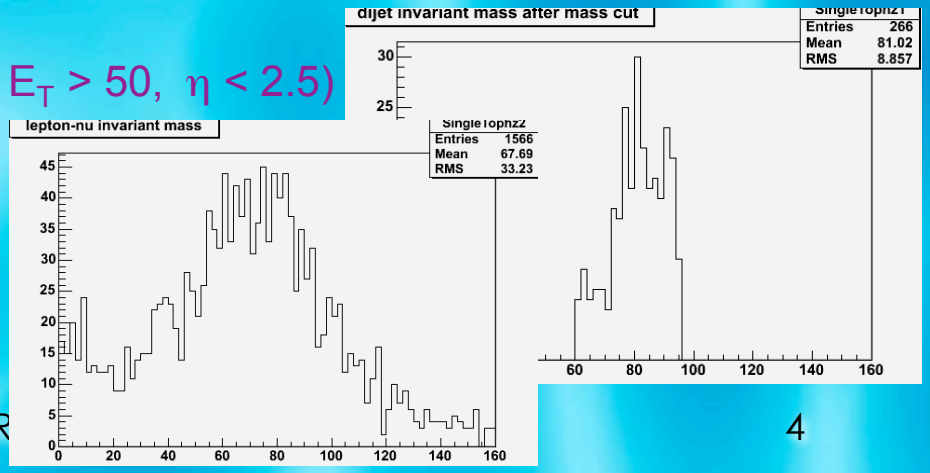
AOD Studies Summary

- 65020 events from rome.004530.recov10.wt_ph_ml.* and rome.004531.recov10.wt_pl_mh.*
- Objects accessed:
  - ElectronCollection
  - METFinal
  - ConeTowerParticleJets (Cone 07)
  - BJetCollection



32028 evts with 1 one P_T ele (XRatio > 0.6)
28582 evts with MET > 20 GeV
12175 evts with 1 and only 1 b-jet (Lhsig > 0.9, E_T > 50, η < 2.5)
1566 evts with 2 jets (3 total) E_T > 30, η < 2.5

2.4% final acceptance (3% TDR)



Goals

VALIDATION:
- We want to arrive to a systematic comparison of CBNT and AOD for fast and full simulation using the Wt channel

To Do List:
-- Ele ID check (IsEM vs Xratio vs Likelihood)
-- B-tagging Efficiency: Standard Algorithms vs Combined Likelihood
-- Adding Muons (an entirely different beast..)
-- Study of jet linearity and energy resolution systematics
-- Full Comparison with TDR and coherence between atlfast and AOD analysis
-- Complete background picture (where are W + jets?)
-- AOB

PHYSICS

-- Benchmark the channel and identify the analysis strategy
-- Understand possible sensitivity to new physics

B-Tagging Studies

Sample

- 20000 events from rome.004531.recov10.wt_pl_mh.*
- Objects accessed:
  - ConeTowerParticleJets (Cone 07)
  - BJetCollection (btagging was run only for cone 0.7 jets)
  - Cone04TowerParticleJets (Cone 0.4)
  - BJetCollection - Btagging was rerun following the instructions at: https://uimon.cern.ch/twiki/bin/view/Atlas/BTagging#Running_the_b_Tagging

Outlook

- Preliminary look at b-tagging efficiency and light jet rejection
- Using as reference the talks of:
  - L. Vacavant, Rome Workshop
  - J.B. deVivie, May 2005 b-tagging group
  - L.Vacavant, Feb 2006, pg15
- In Rome preliminary results, LHSig was used to select b-jets

Summary on b-tagging algs

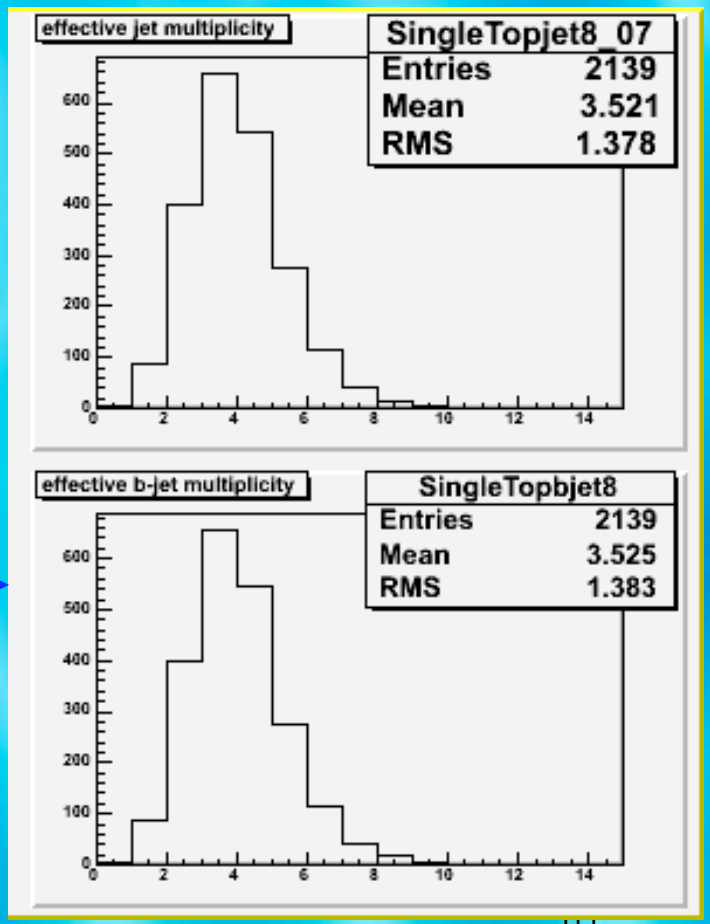
L. Vacavant, Rome workshop

- Historical » taggers:
  - **IP2D**: transverse impact parameter
  - **IP3D**: 2D+longitudinal
  - **SV1, SV2**: inclusive secondary vertex **SV1+IP3D** (called SV1 in CBNT)
- New taggers:
  - Lifetime2D: transverse impact parameter
  - **IhSig**: secondary vertex + impact parameter (2D&3D)
- Tagging weight:
  - IP2D: based on impact parameter significances $S=d_0/\sigma(d_0)$
  - Track weight: likelihood ratio $w_t=P_b(S)/P_u(S)$
  - Jet weight: $W_j= \Sigma \ln w_t^i$
- Generalization of the weight for other taggers, can be combined

by summing them up.

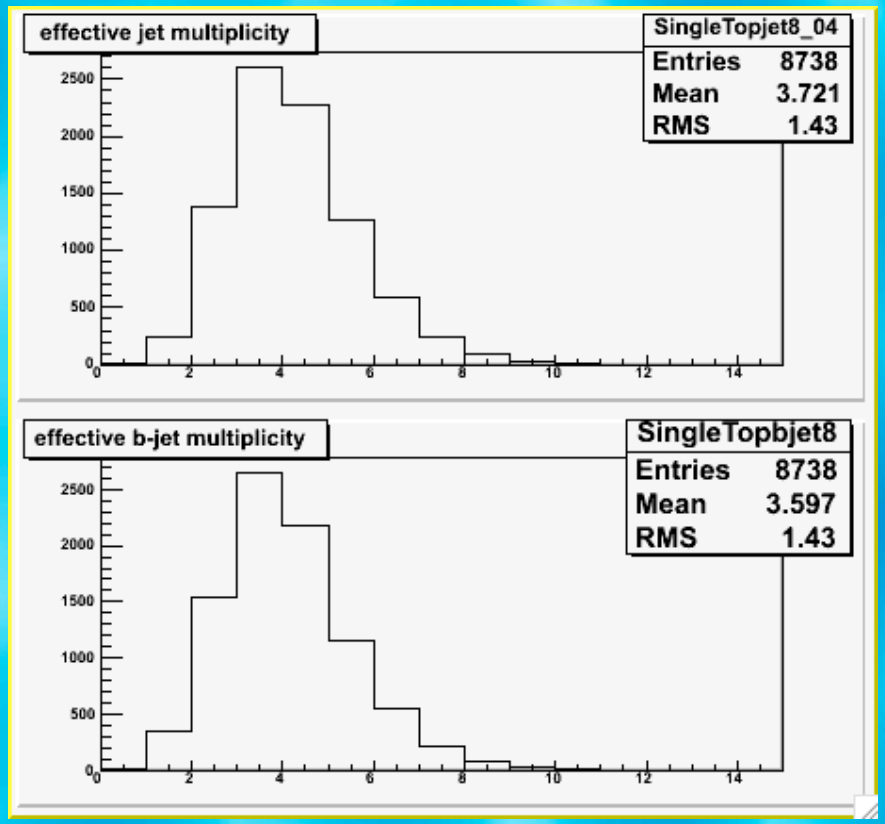
BTagCollection

Btag collection, in Rome samples,
includes only cone07 Jets,
tagged or untagged
(same multiplicity as the
ConeTowerCollection)



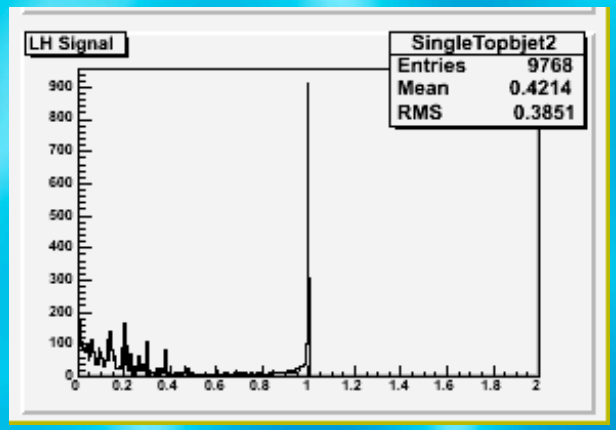
Cone 0.4

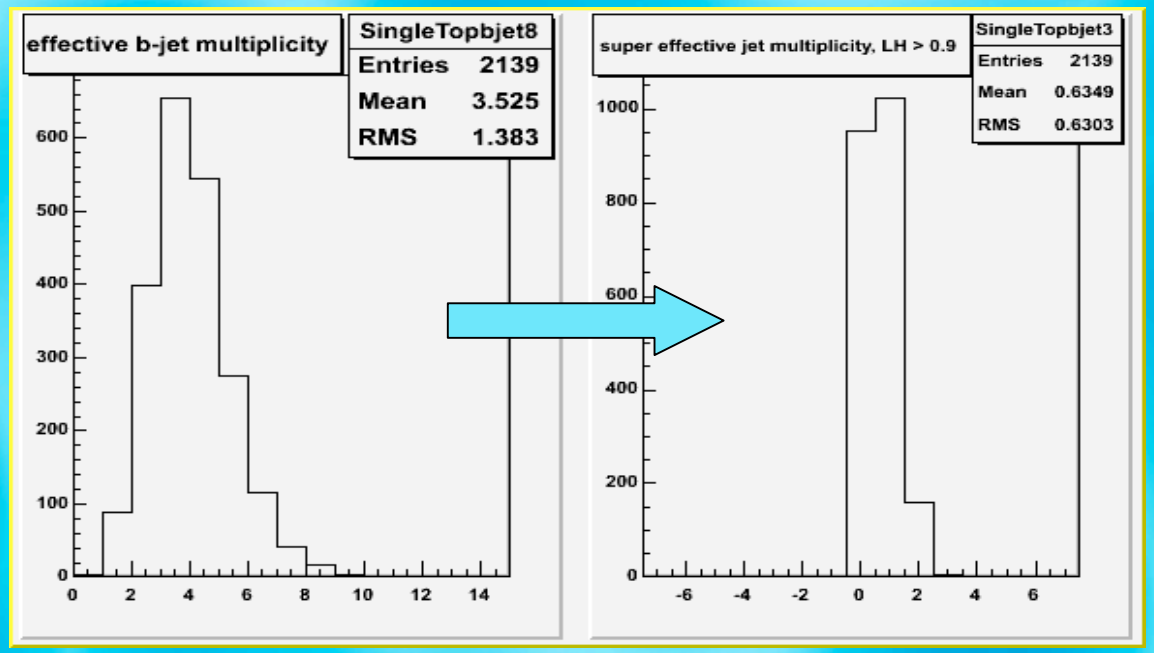
We reprocessed the data as from the recipe on the btagging page
and got the multiplicities for cone 0.4 jets.



Rome selection

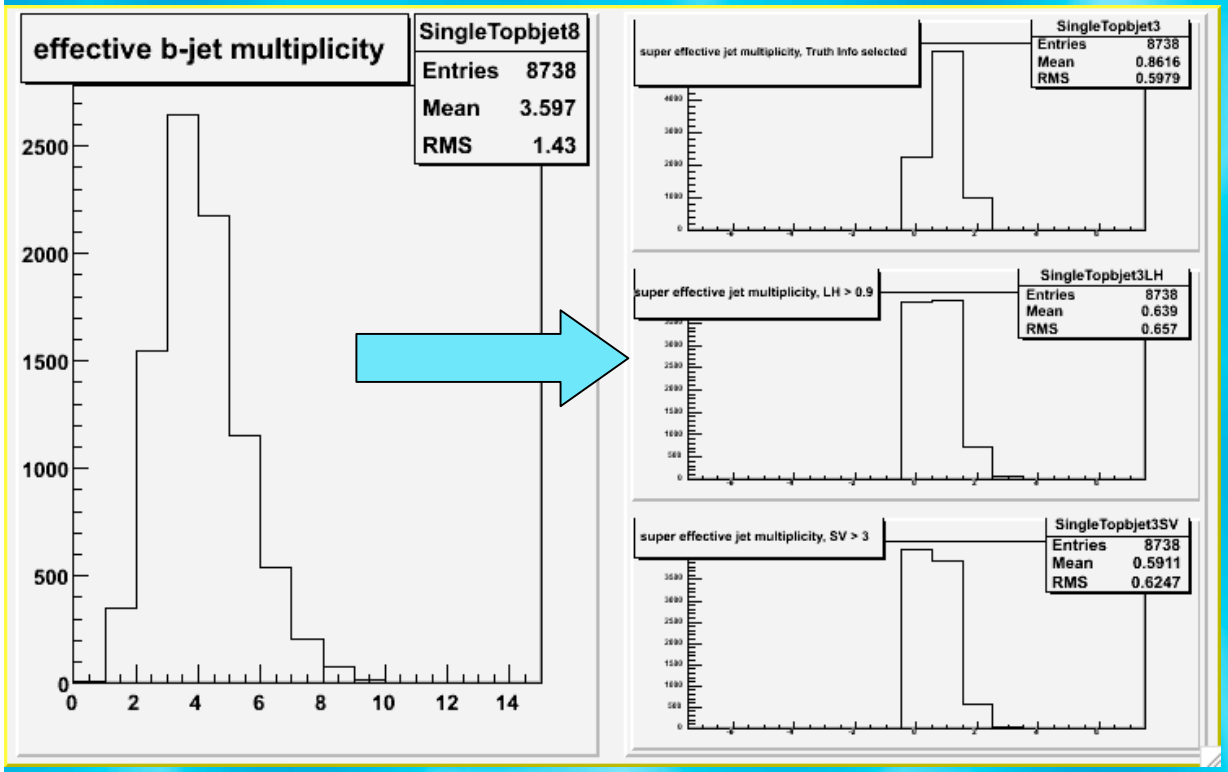
- In the BTagCollection a jet was selected if:
  - $E_T > 50$ GeV, $\eta < 2.5$
  - LHSig > 0.9

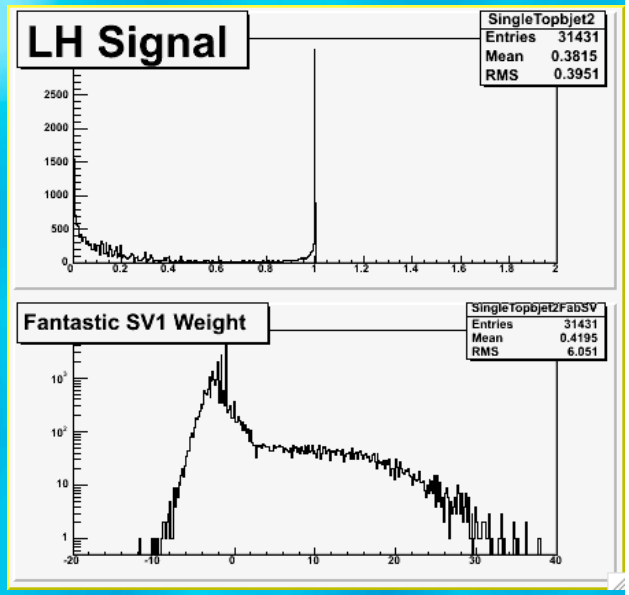




B-jet selection

From the Btag collection jets were selected using TruthInfo, LHSig and SV1





After Rome

- Suggestion to use SV1, IP2D and IP3D
- Weights accessed from AOD:
  - M_bjetwSV1[j] = (*newBJets)[j]->weightForTag("SV1");
  - m_bjetwIP2D[j] = (*newBJets)[j]->weightForTag("IP2D");
  - m_bjetwIP3D[j] = (*newBJets)[j]->weightForTag("IP3D");

- Various web pages/instructions suggest a cut at

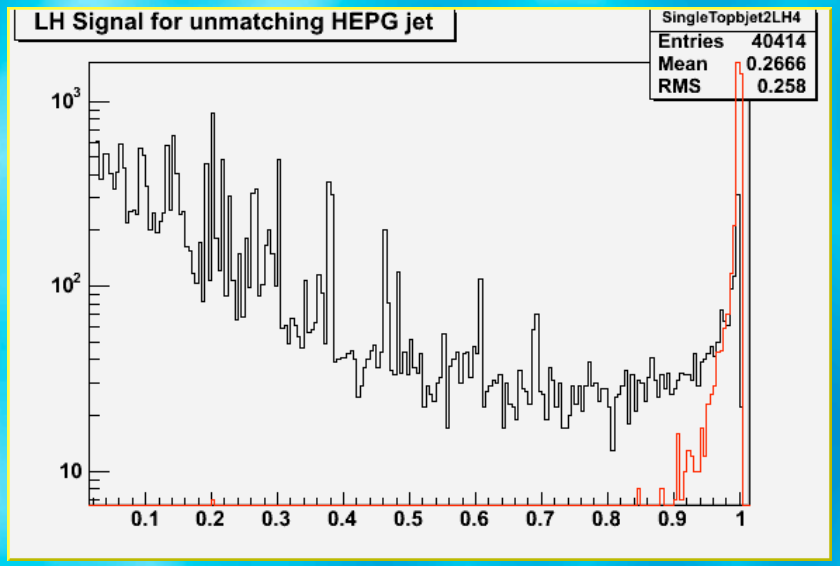
  Weight > 3.0 to select b-jets

- We tested various value of the cut , from 1 to 9 and compared with lhSig.

B-tag efficiencies

LHSig (cut at 0.9) is more efficient than the other algorithms (cut at 3.0)

LHSig distribution:
IP2D > 3.0 (red)
IP2D < 1.0 (black)



Suggestion to re-run the tagging algs, as there were changes After Rome samples were produced.

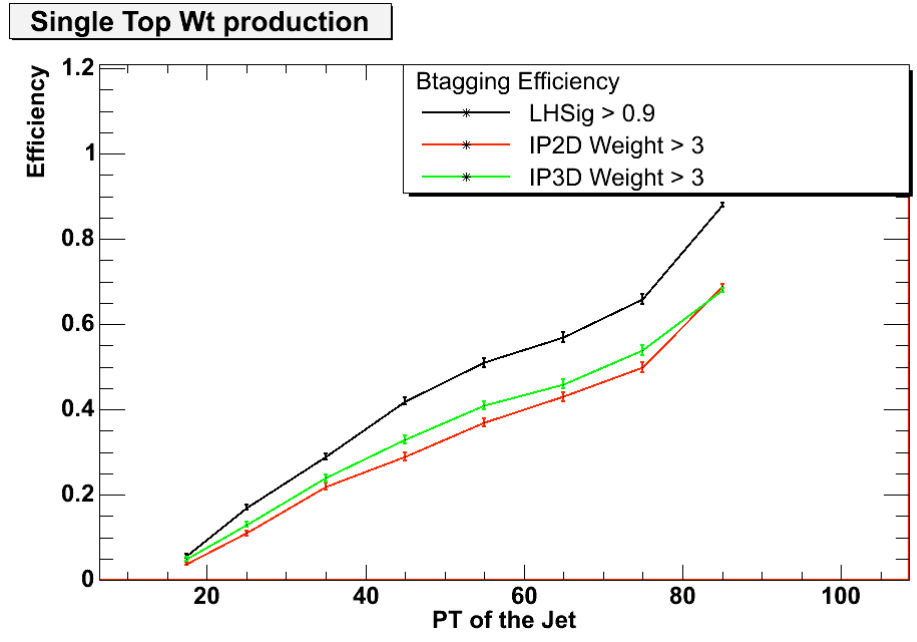
B-tag efficiencies

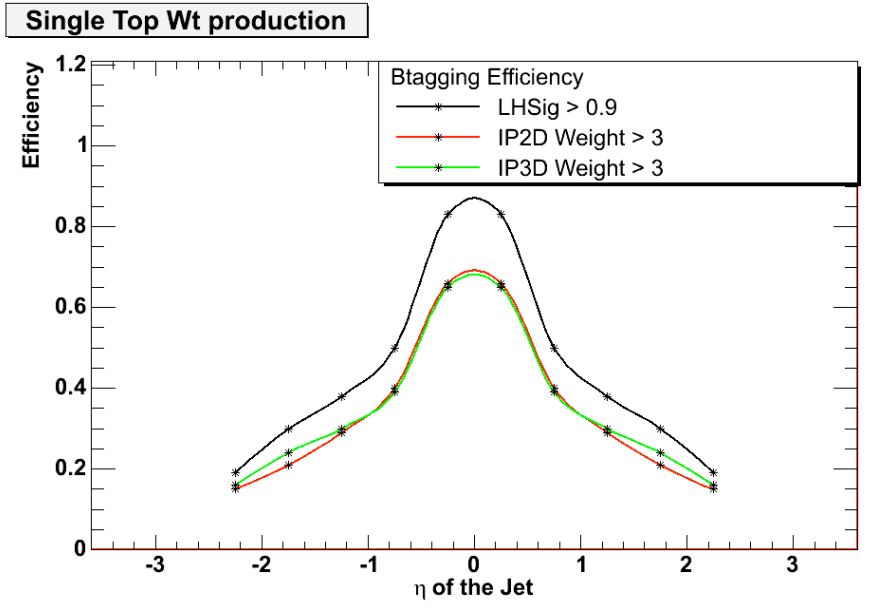
Efficiencies are calculated in the following way:
Denominator: number of jets matched with the b-parton, with $P_T > 50$ GeV, $\eta < 2.5$
Numerator: ditto with cut on weight/likelihood

IP2D Cut	Eff Ip2D	SV1 Cut	Eff SV1	LHSig cut	Eff LHsig
1	0.60 0.63	1	0.63 0.63	0.1	0.80 0.75
2	0.54 0.55	2	0.59 0.59	0.2	0.76 0.72
3	0.49 0.48	3	0.55 0.57	0.3	0.72 0.69
4	0.43 0.41	4	0.53 0.54	0.4	0.70 0.67
5	0.38 0.35	5	0.51 0.51	0.5	0.68 0.66
6	0.33 0.28	6	0.48 0.48	0.6	0.67 0.65
7	0.29 0.21	7	0.46 0.46	0.7	0.65 0.63
8	0.25 0.18	8	0.43 0.43	0.8	0.63 0.61
9	0.21 0.14	9	0.41 0.40	0.9	0.60 0.57

Numbers from Dec 2005 presentation

B-tag efficiencies





Light Jet rejection

In order to reproduce the procedure outlined in Laurent's talk one needs to access the parton level information of the light jets.
This was not done in September (Truth Info missing from our ntuples) when we used an alternative selection using LHSig for both b and light jets.

We updated the results using TruthInfo in December and now we are presenting the results for jets of cone 0.4

b-tagging performance estimators

- b-jet efficiency ε_b:
  - Denominator:
    - jets defined as b using MC truth with (raw) p_T>15 GeV/c, $|\eta|$<2.5
    - NB: jets with no "good" tracks for b-tagging **are** included
    - NB: iso. electrons are not present in the JetTag collection (.)
  - Numerator:
    - ditto + cut on a tagging weight
- light-jet rejection $R_U = 1 / \varepsilon_U$
  - R=100 means 1% mistag rate
  - light jets: u, d, s, g

Light Weight rejection

	R_u (ε_b = 50%)	R_u (ε_b = 60%)
IP2D	166 (125) (158 -109)	25 (50) (55-57)
LHSig	NA (172-NA)	33 (33) (66-NA)
SV1	333 (100) (505-325)	100 (33) (184-156)

Previous presentation

WH sample (L.V.)

ttbar sample (L.V.)

Wt (S.R)

Conclusions

- B-Tag studies on Wt samples:
  - Preliminary tests on various b-tag algorithms, as out of the box on Rome samples for single top were performed
  - Reprocessing of data to obtain cone 0.4 bjets was done;
  - Generally good agreement with previous studies (L.V.)
  - LHSig has higher efficiency to select b-jets (LHSig > 0.9) in Wt data but has a very poor rejection factor.
  - SV1 has slightly lower efficiency, but much higher rejection factor.
  - More studies will be done.
  - More testing woth DC3 data.
  - Planning on a presentation at the btag group sometime in the future